

AMENDMENTS TO THE CLAIMS

1-4. (Canceled)

5. **(Currently Amended)** A videoconferencing system comprising:
a conference bridge for interconnecting a plurality of remotely located
videoconferencing stations; and
a speaker identification subsystem to determine whether a conferee is speaking
based, at least in part, on ~~whether lip movements ascertained~~ distinguishing a
conferee's lips and lip movements from other image features in a digital video signal
from a conference station at which the conferee is located and determining whether the
lip movements are reasonably consistent with an audio signal from the conference
station, the subsystem also to determine which of a plurality of conferees is speaking the
loudest when multiple conferees are speaking simultaneously from different conference
stations.

6. (Previously Presented) The videoconference system of claim 5 wherein the
speaker identification subsystem comprises a voice activity detector.

7-8. (Canceled)

9. (Previously Presented) The videoconference system of claim 6 wherein said voice
activity detector includes image analysis and recognition software.

10. (Previously Presented) The videoconference system of claim 29 wherein said means for visually altering said image comprises means for highlighting a border around said image of said conferee determined to be the loudest speaker.

11. **(Currently Amended)** A videoconference station comprising:

a transmitter to transmit a combined digital audio video signal to a videoconference bridge; and

a speaker identification subsystem located at said videoconference station to determine whether the conferee at said videoconference station is speaking by ~~analyzing whether lip movements of said conferee ascertained from a~~ distinguishing a conferee's lips and lip movements from other image features in a digital video signal at said station and determining whether the lip movements are substantially consistent with an audio signal at said station so as to indicate human speech, the subsystem also to determine which of a plurality of conferees is the loudest speaker when multiple conferees are speaking simultaneously from different videoconference stations.

12. (Previously Presented) The videoconference station of claim 11 wherein said speaker identification subsystem comprises a voice activity detector.

13. (Previously Presented) The videoconference system of claim 12 wherein said voice activity detector includes image analysis and recognition software.

14. (Currently Amended) A method of displaying images of a plurality of conferees in a videoconference system, comprising:

determining whether a conferee is speaking by distinguishing a conferee's lips and lip movements from other image features in a digital video signal from a conference station at which the conferee is located and analyzing a consistency between ascertained lip movements of said conferee ~~ascertained from a video signal from a conference station at which the conferee is located~~ and an audio signal from the conference station such that the combination of lip movement and audio signal indicates human speech;

determining which of the conferees is the loudest speaker when multiple conferees are speaking simultaneously from different conference stations; and

visually altering an image of said conferee that is displayed to other conferees when said conferee is determined to be the loudest speaker.

15-28. (Canceled)

29. (Previously Presented) The videoconferencing system of claim 5, further comprising:

means for visually altering an image of said conferee displayed in other conference stations if said conferee is determined to be the loudest speaker of the plurality of conferees.

30. (Currently Amended) A videoconferencing system comprising:

a speaker identification subsystem located at each of a plurality of remotely located videoconferencing stations, the subsystem to determine whether a conferee is speaking based, at least in part, on ~~whether lip movements of the conferee ascertained from a~~ **distinguishing a conferee's lips and lip movements from other image features in a digital** video signal from a videoconference station at which the conferee is located **and determining whether any ascertained lip movements** are reasonably consistent with an audio signal from the conference station, ~~and~~ **the subsystem** also to determine which of the conferees is speaking the loudest when multiple conferees are speaking simultaneously; and

a conference bridge coupled to each of the videoconferencing stations to visually alter an image of a conferee who is speaking before combining the image with images received from each of the other videoconferencing stations and transmitting a combined image back to each of the videoconferencing stations.

31-34. (Canceled)